In the Description

Please insert the following title to replace the title on page 1 of the English translation:

OPTIMIS~~Z~~ED PROTEIN SYNTHESIS

On page 1, after the title of the invention, please add the following section heading and accompanying paragraph:

CROSS-REFERENCE TO RELATED APPLICATIONS

This application is a U.S. national counterpart application of international application serial no. PCT/EP2003/013964 filed December 9, 2003, which claims priority to German application serial no. 102 57 479.0 filed December 9, 2002.

Please insert the following heading and paragraph to replace the first heading and first paragraph on page 1 of the English translation:

FIELD OF INVENTION ~~Description~~

The invention concerns a method for the optimis~~z~~ed production of proteins in an *in vitro* or *in vitro* expression system and reagents suitable therefor.

Please add the following heading and insert the paragraph below to replace the second paragraph on page 1 of the English translation:

BACKGROUND

Hannig, G. & Makrides, S.C. (1998) Tibtech Vol 16, pp 54-60 have described strategies for optimis~~z~~ing heterologous protein expression in *E. coli* ~~E. coli~~. A key factor in this connection is the efficiency of the initiation of translation in which the usage of particular codons

plays a certain role. Thus, George et al. (1985) DNA Vol 4, pp 273-281, show that the expression of a heterologous gene can be increased by using codons in the region after the start codon that are frequently utilized in *E. coli* ~~E. coli~~ genes. It is predominantly structural elements at the 5' end of mRNA that are particularly important for translation initiation. Makrides (1996) Microbiol. Rev. Vol 60, pp 512-538 described various translation enhancer sequences such as the sequence from the T7 phage gene 10 leader and a U-rich sequence from the 5'-untranslated region of some mRNAs such as the atpE gene of *E. coli* ~~E. coli~~.

Please insert the following paragraph to replace the third paragraph on page 1 of the English translation:

No translation initiation sequences have been described up to now which can be used universally. However, strategies have been described that reduce the potential for the formation of secondary structures at the 5'-end of the mRNA. In particular the ribosomal binding site was enriched with adenine and thymine building blocks. Stenstöm et al. (2001) Gene Vol. 263, pp 273-284 showed that strongly expressed *E. coli* ~~E. coli~~ genes have a high content of adenines especially in the + 2 codon following the start codon. However, there are also many positive and negative exceptions to this rule.

Please insert the following paragraph to replace the first paragraph on page 4 of the English translation:

Hence there is a need to provide a method for the optimis~~z~~ed production of proteins in which the disadvantages of the prior art are at least partially eliminated.

Please add the following heading above the second paragraph on page 4 of the English translation:

SUMMARY

Please add the following three paragraphs between the second and third paragraphs on page 4 of the English translation:

Another subject matter of the invention is a reagent for producing a protein comprising:

(a) a nucleic acid sequence that is heterologous to the nucleic acid sequence coding for the desired protein which can be inserted into the protein-coding nucleic acid sequence in the correct reading frame and which can form a stem-loop structure at a distance of 6-30 nucleotides on the 3' side of the translation start codon, and

(b) an expression system that is suitable for producing the protein.

A method for producing a protein is also provided. The method comprises the steps of (a) providing a nucleic acid sequence coding for the protein wherein the nucleic acid sequence coding for the protein comprises a translation start codon; (b) inserting a heterologous nucleic acid sequence on the 3' side of the translation start codon in the correct reading frame, wherein said heterologous nucleic acid sequence forms a stem-loop structure on the 3' side of the translation start codon 6-30 nucleotides from the 3' side of the start codon; (c) providing an expression system for the protein; (d) introducing the nucleic acid sequences combined in step (b) into the expression system; and (e) forming the stem-loop structure wherein the length of the stem is in the range of 4-12 nucleotides.

A composition for producing a protein is also provided. The composition comprises (a) a nucleic acid sequence that is heterologous to the nucleic acid sequence coding for the protein wherein the heterologous nucleic acid sequence is inserted into the protein-coding nucleic acid sequence in the correct reading frame and wherein the heterologous nucleic acid sequence forms a stem-loop structure 6-30 nucleotides from the 3' side of the translation start codon; and (b) an expression system for the protein.

Please add the following heading above the third paragraph and insert the following paragraph to replace the third paragraph on page 4 of the English translation:

DETAILED DESCRIPTION

The solution according to the invention for a universally optimis~~z~~ed expression construct comprises the insertion of a small heterologous DNA sequence element having preferably a maximum of 201 base pairs, particularly preferably a maximum of 45 base pairs, directly after the start codon of the gene to be expressed, which substantially prevents the formation of stable stem-loop structures in the region of the Shine-Dalgarno sequence and of the start codon and thus results in an optimis~~z~~ed translation initiation and optimis~~z~~ed protein synthesis. Hence a fusion protein is formed in which preferably only a small peptide having a maximum of 67 amino acids and particularly preferably a maximum of 15 amino acids is attached to the desired protein.

Please insert the following paragraph to replace the second paragraph on page 6 of the English translation:

Alternatively the protein can be produced in an *in vivo* expression system in which case it is possible to use a prokaryotic cell e.g. a gram-negative prokaryotic host cell in particular an *E. coli* ~~E. coli~~ cell or a gram-positive prokaryotic cell in particular a *Bacillus subtilis* cell, a eukaryotic host cell e.g. a yeast cell, an insect cell or a vertebrate cell in particular an amphibian, fish, bird or mammalian cell or a non-human eukaryotic host organism as the expression system.

Please add the following paragraph between the third and fourth paragraphs on page 8 of the English translation:

A subject matter of the invention is a method for producing a protein comprising the steps:

(a) providing a nucleic acid sequence coding for the protein in which a heterologous nucleic acid sequence is inserted on the 3' side of the translation start codon in the correct reading frame, said nucleic acid being selected such that a stem-loop structure is formed on the 3' side of the translation start codon at a distance of 6-30 nucleotides,

(b) providing an expression system suitable for expressing the protein and

(c) introducing the nucleic acid sequence according to (a) into the expression system according to (b) under such conditions that the protein is synthesized.

Please add the following heading above the fifth paragraph on page 9 of the English translation:

BRIEF DESCRIPTION OF THE DRAWINGS

Please insert the following paragraph to replace the fifth paragraph starting with "Figure 1" on page 9 of the English translation:

Fig. 1 ~~**Figure 1**~~ shows a schematic representation of the nucleic acid sequence elements necessary for carrying out a two-step PCR.

Please insert the following paragraph to replace the sixth paragraph on page 9 of the English translation:

Fig. 2 ~~**Figure 2**~~ shows a schematic representation of stem-loop structures of different lengths in heterologous nucleic acid sequences used for insertion into GFP expression constructs.

Please insert the following paragraph to replace the seventh paragraph on page 9 of the English translation:

Fig. 3 ~~**Figure 3**~~ shows an evaluation of the results of the expression of GFP using the hairpin-loop GFP constructs of Fig. ~~figure~~ 3 in an RTS expression system. 1 µl of each preparation (duplicate determinations) was separated electrophoretically by SDS-PAGE and blotted on a PVDF membrane. Detection was by means of a DCP Star and Lumi-Imager.

Please insert the following paragraph to replace the first paragraph on page 10 of the English translation:

Fig. 4 ~~**Figure 4**~~ shows a schematic representation of stem-loop structures at different positions in heterologous nucleic acid sequences used to insert GFP expression constructs.

Please insert the following paragraph to replace the second paragraph on page 10 of the English translation:

Fig. 5 ~~**Figure 5**~~ shows the expression of GFP using the heterologous nucleic acid sequences shown in Fig. ~~figure~~ 4. The experiments were carried out and evaluated as described in the legend to Fig. ~~figure~~ 3.

Please insert the following paragraph to replace the third paragraph on page 10 of the English translation:

Fig. 6 ~~Figure 6~~ shows an evaluation of the results of the expression of the CIITA gene (wild-type: lane 1; mutants lanes 2-10) using different heterologous nucleic acid sequences with stem-loop structures.

Please insert the following paragraph to replace the fourth paragraph on page 10 of the English translation:

Fig. 7 ~~Figure 7~~ shows an evaluation of the results of the expression of the CMV capsid (1049) gene (wild-type: lane 1; mutants lanes 2-10) using different heterologous nucleic acid sequences with stem-loop structures.

Please insert the following paragraph to replace the fifth paragraph on page 10 of the English translation:

Fig. 8 ~~Figure 8~~ shows an evaluation of the results of the expression of the survivin gene (wild-type: lane 10; mutants lanes 1-9) using different heterologous nucleic acid sequences with stem-loop structures.

Please insert the following paragraph to replace the sixth paragraph on page 10 of the English translation:

Fig. 9 ~~Figure 9~~ shows an evaluation of the results of the expression of the GFP gene (wild-type: lane 10; mutants lanes 1-9) using different heterologous nucleic acid sequences with stem-loop structures.

Please insert the following paragraph to replace the seventh paragraph on page 10 of the English translation:

Fig. 10 ~~Figure 10~~ shows an evaluation of the results of the expression of the GFP and the 1049 gene using different heterologous nucleic acid sequences with and without stem-loop structures.

Please insert the following paragraph to replace the first paragraph on page 11 of the English translation:

Fig. 11 ~~Figure 11~~ shows an evaluation of the results of the expression of the CIITA and the survivin gene using different heterologous nucleic acid sequences with and without stem-loop structures.

Please insert the following paragraph to replace the second paragraph on page 11 of the English translation:

Fig. 12 ~~Figure 12~~ shows a schematic representation of two different stem-loop structures in the heterologous sequences according to the invention.

Please insert the following paragraph to replace the third paragraph on page 11 of the English translation:

Fig. 13 ~~Figure 13~~ shows an evaluation of the results obtained with the stem-loop structures shown in Fig. ~~figure~~ 12.

Please insert the following paragraph to replace the fourth paragraph on page 11 of the English translation:

Fig. 14 ~~Figure 14~~ shows a representation of the *in vivo* protein expression of RNA stem-loop constructs compared to the wild-type genes in a Western Blot. Expression of three

independent clones of the RNA stem-loop mutants of the CMV capsid protein 1049 (lanes 1 to 3) and of the CMV capsid protein 1049 wild-type (lanes 4 to 6). Expression of independent clones of survivin RNA stem-loop mutants (lanes 7 to 9) and of the survivin wild-type (lanes 10, 11).

Please insert the following paragraph to replace the fifth paragraph on page 11 of the English translation:

A two-step PCR can be used to amplify genes that are to be expressed and to provide them with the appropriate control regions such as the T7 promoter, T7 gene 10 leader (g10), ribosomal binding site (RBS) and T7 terminator. In the first step the gene is amplified by means of a pair of primers (A, B) which are each complementary over a length of 15 bases with the corresponding gene and contain 15 additional bases which are complementary to a second primer pair (C, D). The second primer pair contains all important regulatory elements which are thus attached to the gene in a second PCR amplification (see Fig. ~~figure~~ 1).

Please insert the following paragraph to replace the second paragraph on page 16 of the English translation:

A schematic representation of the mRNA secondary structures of the hairpin loop GFP constructs is shown in Fig. ~~figure~~ 2.

Please insert the following paragraph to replace the third paragraph on page 16 of the English translation:

After expression in the RTS according to example 2 the amount of GFP formed was measured in a fluorimeter for the purposes of verification and the Western Blot was

quantitatively analysed by CDP-Star detection and evaluation in a Lumi-Imager. The results are shown in Fig. ~~figure~~ 3.

Please insert the following paragraph to replace the third paragraph on page 18 of the English translation:

These DNA constructs with the secondary structures shown in Fig. ~~figure~~ 4 were also synthesized by a two-step PCR using the previously described primers B, C and D and used directly from the PCR reaction as templates in expression preparations. It was ensured that the same amounts of template were used by quantification on an agarose gel with the DNA marker VII and evaluation of this gel in a Lumi-Imager. The expression mixtures were evaluated by a Western Blot. The results are shown in Fig. ~~figure~~ 5.

Please insert the following paragraph to replace the second paragraph on page 19 of the English translation:

A heterologous nucleic acid sequence with a hairpin loop and a stem length of 7 bases at a distance of 15 bases after the start codon was introduced for three of these genes, survivin, cytomegalovirus capsid protein 1049 (1049) and Class II transactivator (CIITA). The wild-type gene (see below *) without the start ATG was placed directly after the hairpin loop. AT-rich sequences were placed in front of the hairpin loop which are able to form less stable base pairs than GC-rich sequences. Furthermore, care was taken that no rare codons for *E. coli* ~~E. coli~~ were used within the introduced sequences.

Please insert the following paragraph to replace the second paragraph on page 30 of the English translation:

The expressions shown in Figs. ~~figures~~ 6 to 9 show that DNA templates synthesized with the stem-loop structures in all cases resulted in protein synthesis whereas no protein synthesis took place with the wild-type gene. The expression of mutant 9 with the hexa-histidine sequence is not quite as good as that of the other AT-rich sequences but has the advantage that the protein that is formed can be purified on Ni-NTA chelate columns by means of this six histidine residue label. Even in the case of the GFP gene which is a gene that is in any case expressed well, the stem-loop constructs resulted in an increase in yield.

Please insert the following paragraph to replace the third paragraph on page 31 of the English translation:

For this a new stem-loop (loop') having the sequence CAG.ACA.AAT. AGA.TAT. TTG.TCT.GTA (G = -9.8 kcal/mol and a stem length of 9 base pairs) was combined with the AT-rich sequence of mutant 1 instead of the original stem-loop sequence CTG.CAC.GTG.ATC.GTG.CAG (G = -9.8 kcal/mol and a stem length of 7 base pairs) for the examples survivin, CIITA and 1049. The two structures are shown in Fig. ~~figure~~ 12.

Please insert the following paragraph to replace the second paragraph on page 32 of the English translation:

The expressions in Fig. ~~figure~~ 14 show that the stem-loop constructs for the two examined gene cytomegalovirus capsid protein 1049 as well as survivin also exhibited a substantially higher expression *in vivo* than the wild-type genes. This proves that the results of the *in vitro* expression can also be applied to *in vivo* expression.